BEST AVAILABLE COPY

Exhibit 3
10/735,256

The leucine-rich repeat: a versatile binding motif

Bostjan Kobe and Johann Deisenhofer

Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. All proteins containing these repeats are thought to be involved in protein–protein interactions. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to β–α structural units. These units are arranged so that they form a parallel β-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features may be responsible for the protein-binding functions of proteins containing leucine-rich repeats.

LEUCINE-RICH REPEATS (LRRs) were first discovered in leucine-rich α2-glycoprotein, a protein of unknown function from human serum[1]. LRR-containing proteins now represent a diverse group of molecules with differing functions and cellular locations in a variety of organisms (Table 1, Fig. 1). LRRs are usually present in tandem, and the number of LRR motifs ranges from one, in platelet glycoprotein Ibβ, to 30, in chaoptin. The most common length of an LRR is 24 residues, but repeats containing any number between 20 and 29 residues are found. The recent X-ray structure analysis of ribonuclease inhibitor (RI) from pig liver[2] (Fig. 2), a protein built entirely from LRRs, provided the first picture of the three-dimensional architecture of LRRs.

Consensus sequences of LRRs and evolutionary implications

As the name implies, LRRs are distinguished by a consensus sequence consisting predominantly of leucines (Table I). LRRs are easily overlooked by the commonly used methods for sequence comparison because they contain gaps, vary in length and amino acid composition, and therefore have a degenerate consensus sequence. Consequently, LRRs in many proteins were not detected for some time.

The consensus sequence compiled from all known LRR-containing proteins contains leucines or other aliphatic residues at positions 2, 5, 7, 12, 16, 21 and 24, and asparagine, cysteine or threonine at position 10 (Table I). Most proteins contain exclusively asparagine at position 10, but three have exclusively cysteine in this position (Table I). The functional and evolutionary significance of the variations of consensus residues at position 10 is not known.

Consensus sequences derived from LRRs in individual proteins often contain additional conserved residues in positions other than those mentioned above; mostly these are aliphatic and aromatic amino acids, sometimes glycines and prolines, but seldom other amino acids. The hydrophobic consensus residues in the carboxy-terminal parts of the repeats are commonly spaced by 3, 4 or 7 residues.

About half of the genes encoding the known LRR-containing proteins have been sequenced. Gonadotrophin receptor genes contain several similarly sized exons in their LRR domains that are homologous in their exon–intron junction sequences[3], suggesting that the LRR domain evolved by exon duplication and shuffling from a single prototypic exon corresponding to one LRR. Such an evolutionary relationship is less obvious in genes encoding other LRR-containing proteins, but the introns often localize at similar positions in the repeats (most frequently between positions 1 and 6, less frequently between positions 11 and 18, and rarely elsewhere). The most likely mechanism of evolution of LRR-containing proteins would therefore comprise unequal crossovers and duplications of gene fragments corresponding to prototypic leucine-rich building blocks. Two evolutionary scenarios are possible: (1) an ancestral LRR module evolved first, and duplicated to give rise to individual members of the superfamily that later developed their own unique consensus sequences; or (2) the duplications occurred separately in each protein family and, owing to structural constraints, gave rise to similar, but unique, consensus sequences. These two schemes are not mutually exclusive and there is good reason to believe that the divergent and convergent scenarios have worked in conjunction to yield the LRR superfamily that exists today. The variation in the length of LRRs among proteins and the presence of consensus sequences unique to a particular protein argue for a rich and complex evolutionary history that may have sprung from more than one root.

Three-dimensional structure of LRRs

The recently determined crystal structure of porcine ribonuclease inhibitor protein (RI)[2] (Fig. 2a) has allowed us to examine the three-dimensional architecture of LRRs. RI is a cytoplasmic protein that inhibits ribonucleases from the pancreatic superfamily by binding very tightly to an extensive surface area of ribonuclease containing its catalytic site[4]. The sequence of RI consists of 15 alternating LRRs of 28 (A-type) or 29 (B-type) residues each (Fig. 2b). Ninety percent of the structure is built of these repeat units. The short amino- and carboxy-terminal segments display a weaker homology to the internal repeat motif.

In the RI molecule, LRRs correspond to β–α structural units, consisting of a short β-strand and an α-helix approximately parallel to each other (Fig. 2). All repeats, including the terminal segments, adopt very similar structures. All structural repeats contain 28 or 29 residues, except the amino-terminal repeat, which contains 25 residues. The structural units are arranged so that all the β-strands and the helices are parallel to a common axis, resulting in a nonglobular, horseshoe-shaped molecule with a curved parallel β-sheet lining the inner circumference of the horseshoe, and the helices flanking its outer circumference.

B. Kobe and **J. Deisenhofer** are at the Howard Hughes Medical Institute and Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235-9050, USA; **B. Kobe** is on leave from the Jozef Stefan Institute, Department of Biochemistry and Molecular Biology, Jamova 39, 61111 Ljubljana, Slovenia.

© 1994, Elsevier Science Ltd 0968–0004/94/$07.00

Table I. Proteins with leucine-rich repeats

Protein (species)[a]	Function–ligand[a]	Location[a,b]	Repeats[c]	Length[d]	Consensus sequence[e]	PIR[f] entry
					5 10 15 20 25	
RNase inhibitor (porcine)	RNase inhibitor–RNase	Cytoplasm	15	28 (A)	.LE.L.L..C-.LT...C..L..aL....	A31857
				29 (B)	.L.EL.L..N-.LGD.Ga..L...L..P..	
Leucine-rich α2-GP (human)	?–?	Serum	8	24	.L..L.L..N-.L..-L...LL...--.	NBHUA2
RNA1 (*Saccharomyces cerevisiae*)	RNA processing–?	Cytoplasm	8	29	.L..L.L..N-.a.....a..a.....	BVBYN1
U2 snRNP A' (human)	Splicing–U2 snRNP	Nucleus	4	24	.L..L.a..N-.a..-.....L--.	S03616
Biglycan (human)	ECM binding–laminin, fibronectin, TGF-β	ECM	8	24	.L..L.L..N-.I..-a....a...--.	A40757
Decorin (human)	ECM binding–collagen, fibronectin, thrombospondin, TGF-β	ECM	10	24	.L..L.L..N-.I..-V....a...--.	NBHUC8
Fibromodulin (bovine)	ECM binding–collagen, fibronectin	ECM	11	24	.L..L.L..N-.a..-a....a...--.	S05390
Lumican (chicken)	Corneal transparency–?	ECM	12	24	.L..L.L..N-.L..-.....a...--.	A41748
Proteoglycan-Lb (chicken)	?–?	ECM	6	24	.L..a.L..N-.I..-.....a...--.	A41781
Osteoinductive factor (bovine)	Bone morphogenesis–BMP	ECM	6	24	.L..a.L..N-.a..-.....F....--.	A35272
Platelet GP Ibα (human)	Cell adhesion–vWF, thrombin	PM (EC)	7	24	.L..L.L..N-.L..-LP.GL...L--.	NBHUIA
Platelet GP V (human)	Cell adhesion–GP IX, GP Ib	PM (EC)	14	24	.L..L.L..N-.L..-LP..LF..L--.	–
YopM (*Yersinia pestis*)	Virulence factor–thrombin	IC + EC	12	20	.L..L.a..N-.L..-LP-.----L--PP	A33950
IpaH7.8 (*Shigella flexneri*)	?–?	?	6	20	.L..L.V..N-.L..-LP-.----L--P.	A35149
IpaH4.5 (*Shigella flexneri*)	?–?	?	8	20	.L..L.a..N-.L..-LP-.----L--P.	S18248
Toll (*Drosophila*)	Embryo development–?	PM (EC)	19	24	.L..L.L..N-.L..-.....F...--.	A29943
Slit (*Drosophila*)	Axon development–?	EC	19	24	.L..L.L..N-.I..-.....F..L--.	A36665
Connectin (*Drosophila*)	Synapse development–?	PM (EC)	7	24	.L..LNL..N-.I..-a...aF..L--.	S28484
Chaoptin (*Drosophila*)	Photoreceptor-cell development–?	PM (EC)	30	24	.L..L.L..N-.a..-a....F..a--.	A29944
Flightless-I (*Drosophila*)	Embryo development–?	PM (EC)	16	23	.L..L.LS.N-.L..-aP..a..-L--.	–
Oligodendrocyte myelin GP (human)	Myelination–?	PM (EC)	8	24	.L..L.LSNN-.a..-a.....L--.	A34210
CD14 (human)	Cell-surface receptor–LPS–LPB	PM (EC)	8	27	.a..L.L..N-.....	TDHUM4
Trk (human)	Receptor protein kinase–NGF	PM (EC)	2	23	.L..L.LS.N-.L.....-.-....--.	TVHUTT
TrkB (mouse)	Receptor protein kinase–BDNF, NT-3	PM (EC)	3	23	.L..L.aT.N-.LTS...-.-....--T	S06943
TrkC (porcine)	Receptor protein kinase–NT-3	PM (EC)	3	23	.LR.aNLSQN-.L...S.-.-....--.	A40026
TMK1 (*Arabidopsis thaliana*)	Receptor protein kinase–?	PM (EC)	11	23	.L..a.L..N-...G.aP-.-a.SL--.	JQ1674
LH-CG receptor (rat)	Signal transduction–LH, CG	PM (EC)	5	25	.L..L.a..T-.a.....F...--.	A41343
FSH receptor (rat)	Signal transduction–FSH	PM (EC)	7	25	.L..L.aS.T-.....LP...a..a--.	A34548
TSH receptor (dog)	Signal transduction–TSH	PM (EC)	6	25	.a..L.a.NN..a.S-a....a...--.	A40077
Adenylate cyclase (*Saccharomyces cerevisiae*)	Signal transduction–RAS	PM (cytoplasm)	20	23	.L..L.L..N-.a..-a...-a..L--.	OYBY
T-LR (*Trypanosoma brucei*)	?–?	?	18	23	.L..L.LSGC..a..--a..-a..L--.	A36359
RAD1 (*Saccharomyces cerevisiae*)	DNA repair–RAD10	Nucleus	3	23	.a.LaDI..N--LP..a.....N---.	DDBYD1
RAD7 (*Saccharomyces cerevisiae*)	DNA repair–?	?	5	26	.L..L.a..C..a....a...a...--P	A25226
DRT100 (*Arabidopsis thaliana*)	Recombination–?	Chloroplast	5	24	.L..LNL..N-.L.G.IP.S-a.S.--.	A46260
GRR1 (*Saccharomyces cerevisiae*)	Signal transduction–?	Cytoplasm	9	26	.L..a.L..C.NaTD..a..-L..L.-.	A41529
CCR4 (*Saccharomyces cerevisiae*)	Transcription–?	?	4	23	.L..L.a..N-.LT.-LP.E-a...--.	S31286
sds22 (*Schizosaccharomyces pombe*)	Mitosis–dis2, sds21	Nucleus	11	22	.L..L.a..N-.I..-a--ENa..L--.	A38439
p34 ribosome-binding protein (rat)	RM membranes–ribosome	RM membrane (cytoplasm)	4	24	.L..LDL..N-.L..-LP...F..L--.	–
Carboxypeptidase N (human)	Stabilization–catalytic subunit	Plasma	12	24	.L..L.L..N-.L..-LP..aP..L--.	A34901
Internalin (*Listeria monocytogenes*)	Invasion–?	Cell wall	13	22	NL..L.L..N-QISDI.P---L..L--T	A39930
InlB (*Listeria monocytogenes*)	?–?	?	6	22	.L..L.L..N-.L.DI..---L..L--.	C39930
					5 10 15 20 25	
LRR superfamily					.L..L.L..N±.a..±a±±±±a±±a±±.±±	

[a]A single example is given for proteins where sequences from several species are known. Proteins with less than two LRRs are not shown. Abbreviations: RNase, ribonuclease; GP, glycoprotein; snRNP, small nuclear ribonucleoprotein particle; ECM, extracellular matrix; PM, plasma membrane; EC, extracellular; TGF, transforming growth factor; IC, intracellular; BMP, bone-morphogenic protein; vWF, von Willebrand factor; LPS–LPB, complex of lipopolysaccharide and lipopolysaccharide-binding protein; NGF, nerve growth factor; BDNF, brain-derived neurotrophic factor; NT-3, neurotrophin-3; LH, lutrophin; CG, choriogonadotrophin; FSH, follitrophin; TSH, thyrotrophin; T-LR, trypanosomal leucine-rich protein; RM membrane, rough microsomal membrane.

[b]Cellular distribution of the protein. For membrane-associated proteins, the location of the LRR domain is indicated in parentheses.

[c]Total number of repeats is the number of occurrences of the a..a.a..N/C/T sequence, where 'a' represents A, V, L, I, F, Y or M; repeats shorter than 18 residues and isolated single repeats were not counted. Only the counted repeats were used to determine the consensus sequence. Because other authors may have used different criteria, the numbers of repeats, the lengths and the consensus sequences may vary from those originally reported.

[d]The most common length of all the repeats in the protein is given. For some proteins, there are significant deviations from this number. RNase inhibitor is aligned so that two types of repeats (A-type and B-type) alternate in the sequence.

[e]The numbers above the sequences indicate the position in the repeat in reference to the consensus of porcine RNase inhibitor. One-letter code is used for amino acids. An amino acid is included in the consensus if present at that position in more than half of the repeats; 'a' represents A, V, L, I, F, Y or M, and is included in the consensus if these amino acids are present at that position in more than 80% of the repeats. Symbols used: '.', any amino acid; '-', gap; '±', amino acid may or may not be present at this position.

[f]Protein Identification Resource sequence data bank. References for the proteins without PIR entries: Platelet GV, Hickey, M. J. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 8327–8331; Flightless-I, Campbell, H. D. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 11386–11390; p34 ribosome-binding protein, Ohsumi, T. *et al.* (1993) *Biochem. J.* 294, 465–472.

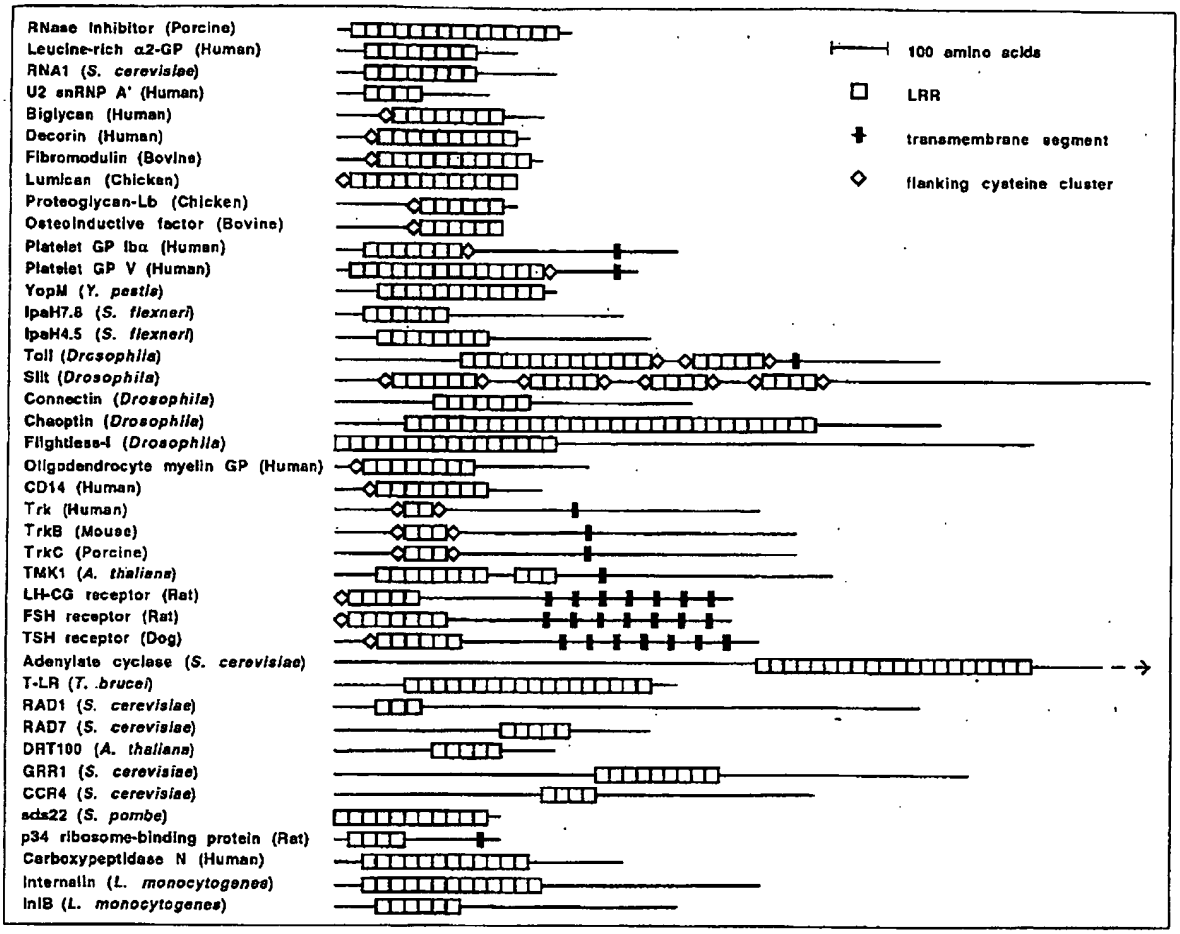

Figure 1
Schematic diagrams of the primary structures of proteins with leucine-rich repeats. For abbreviations and references, see Table I.

The consensus residues of the repeats play structural roles. The sidechains of leucines and of other aliphatic residues (positions 2, 5, 7, 12, 17, 20, 23 and 24), of cysteines (position 17), and of asparagines or cysteines (position 10) form the core of the protein [note that the consensus positions specific to RI are slightly different from the overall consensus of LRR-containing proteins (Table I)]. The spacing of leucines is responsible for proper packing of the β-strands and α-helices. The remaining consensus residues in RI have other roles. For example, in the B-type repeat, glycine at position 13 of the repeat seems to be a helix-initiation signal, while proline at position 27 is a helix-termination signal. Because of the arrangement of the β–α units, sidechains of consensus residues align in stacks, both at the core and on the surface of the protein.

It is unlikely that a single leucine-rich motif would adopt the conformation observed in the RI protein. Neighboring repeats form many polar and nonpolar contacts; also, the β-strands tend to pack against the helices that precede and follow the β-strand in question. Synthetic peptides that correspond to a single repeat have low propensity to adopt a defined structure[5,6].

Secondary-structure predictions of LRRs show no clearly repetitive patterns[1]. Consequently, before the structure of RI had been determined, both an α-helix[7] and a β-strand conformation[8] was proposed as the major secondary structural element of LRRs in different proteins. β-turns were, however, correctly predicted in the vicinity of the conserved asparagine[9].

Leucine-rich repeats should not be confused with the more widely known leucine zippers. Leucine zippers are motifs found in oligomeric proteins, including many DNA-binding proteins such as the products of the c-*fos* and c-*jun* proto-oncogenes. Leucine zippers are built of heptad repeats with leucines present at every seventh position. These leucines form a hydrophobic ridge on one side of an α-helix that interacts with a complementary helix[10]. The only similarity between LRRs and leucine zippers is the conservation of leucines at particular spacings; these leucines cannot usually be replaced by other hydrophobic amino acids. However, the spacing between leucines differs between LRRs and leucine zippers, with only certain leucines in the carboxy-terminal part of LRRs spaced by seven residues. The function of leucines in LRRs and leucine zippers is also different; while leucines in leucine zippers participate in oligomerization, the buried leucines in

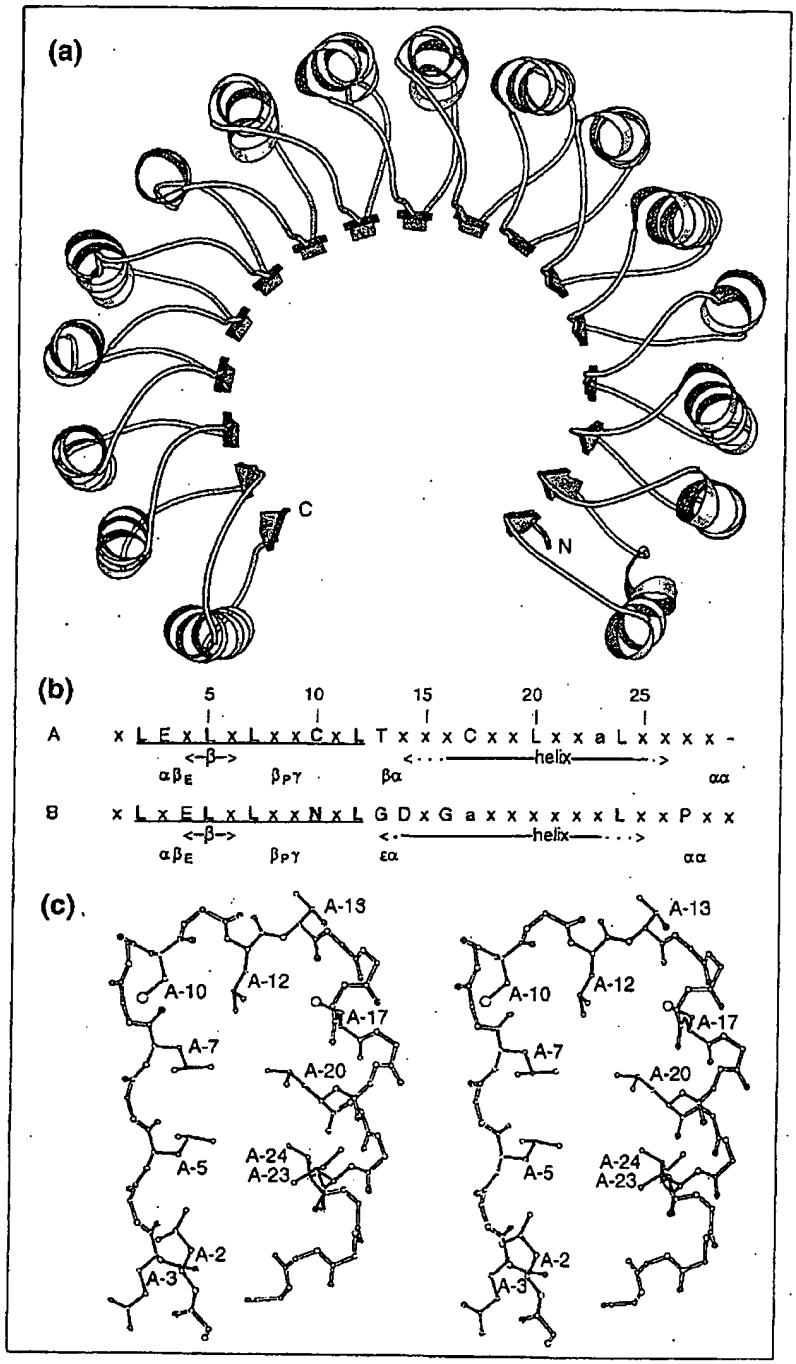

Figure 2

Structure of ribonuclease inhibitor (RI). **(a)** Ribbon diagram of the structure of porcine RI. **(b)** Consensus sequences and secondary structure of the A- and B-type leucine-rich repeats (LRRs) of porcine RI. One-letter amino acid code is used; 'x' indicates any amino acid, 'a' denotes an aliphatic amino acid. The part of the repeat that is strongly conserved throughout the LRR superfamily is underlined, and the conserved residues are bold. Below the sequence, elements of secondary structure are indicated. The solid line marks the core region of the β-sheet or α-helix; the dots denote extensions of the helix in different repeats. The types of β-turns are shown below the corresponding i+1 and i+2 positions. **(c)** Stereo plot of the main chain of a representative leucine-rich repeat corresponding to residues 195–222 of porcine RI. Sidechains of only the consensus residues are shown; bonds in the conserved residues are filled.

LRRs play a structural role and cannot be directly involved in protein–protein interactions.

RI as a model for other LRR-containing proteins

We believe that the repeats in RI are a satisfactory structural model for LRRs in other proteins. Most consensus residues that play key structural roles are conserved throughout the superfamily, especially the consensus residues that cluster around the β-strand region (Table I, Fig. 2). Although the LRRs in RI are among the longest found in the superfamily, shorter repeats could easily adopt a similar β–α conformation, as demonstrated by the amino-terminal structural repeat in RI, which contains 25 residues.

We must, however, allow the possibility that in certain LRR-containing proteins the helical parts might be substituted with more extended structures. This would most probably occur in those proteins that contain shorter variants of LRRs which, in the carboxy-terminal regions of the repeats, lack the appropriate pattern of hydrophobic residues expected for an amphipathic helix, and contain prolines and other amino acids with low propensities to form α-helices.

Substitution of the helical part of the repeats with a more extended structure would influence the overall shape of the protein considerably, particularly its curvature. Two protein folds exist that could shed light on the plausible architecture of a structure with β-strands substituting for α-helix. These are the β-roll folds of pectate lyase C[11] (PelC) and alkaline protease[12] (AP) (Fig. 3). Both β-roll folds are assembled from repetitive structural units and expose a surface of a parallel β-sheet to solvent. In addition, in PelC, as in RI, a ladder of asparagine residues at the carboxy-terminal ends of the β-strands hydrogen bonds to the backbone. The length of a structural repeat of the PelC-type 'β-helix' is 22 residues, and the length of a turn of the AP-type β-roll is 18 residues. Neither PelC nor AP contain LRRs.

In proteins that contain over 21 LRRs and retain the curvature of the structure of RI, the overall organization of the repeats must be different from that in RI to prevent the termini from colliding. Such a collision could be avoided by a twist in the sheet. Alternatively, smaller numbers of repeats could constitute separate domains, linked by segments adopting conformations different from the β–α units.

LRR-containing proteins: functional and evolutionary similarities

LRRs are found in a functionally and evolutionarily diverse set of proteins. Although few properties are shared among all members of the LRR superfamily, two main themes emerge: all LRR-containing proteins appear to be involved in protein–protein interactions, and at least half of them take part in signal transduction pathways (Table 1). LRRs must therefore provide a structural skeleton for the achievement of diverse, specific molecular interactions.

A property of the structure of the LRR that may be helpful in protein binding is its nonglobular shape. Because of this shape, the area available for interaction with smaller globular proteins is substantially increased, facilitating more interactions, and thus enhancing affinity.

Another property of the structure of LRRs that might be useful for protein binding is the exposed surface of the parallel β-sheet. This structural feature is rare in proteins, and has only been discovered recently in structures of PelC[11], AP[12] and RI[2]. The β-roll structures of PelC and AP were proposed to be responsible for transport across membranes. However, such a surface feature could also aid in protein binding. Parallel β-sheets are common in the interior of proteins, where helices or other parallel β-sheets pack against them. In the case of LRR-containing proteins, the interaction with the ligand could substitute for such packing interactions.

In addition to the properties common to most LRR-containing proteins, such as involvement in protein–protein interactions and signal transduction, further similarities restricted to fewer members of the superfamily can be found, on the basis of both their functions and the characteristics of their repeats. These often point to evolutionary and functional relationships among LRR-containing proteins. The specificity of protein–protein interactions of LRR-containing proteins and, therefore, the basis of their functions are most probably due to the specific composition of nonconsensus residues, and might also be influenced by the lengths of the repeats and the flanking domains. These functional and evolutionary relationships are discussed below, and the remarkable diversity of functions of LRR-containing proteins is demonstrated.

Relatives of RI. The closest known relatives of RI seem to be the RNA1 protein, which is involved in RNA processing[13], and the A' protein from the U2 small nuclear ribonucleoprotein particle (snRNP), which is required for snRNP assembly[8]. RI and RNA1 also have the most closely related patterns of conserved residues among the LRR proteins. RI and A' protein both interact with RNA-binding proteins, and therefore may be functional relatives.

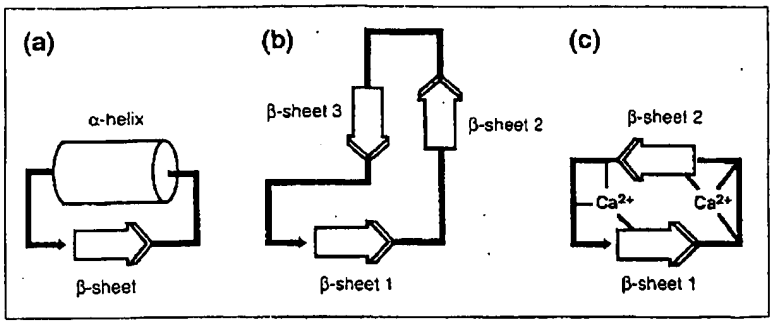

Figure 3
Schematic diagrams of the repetitive structural units of **(a)** ribonuclease inhibitor, **(b)** pectate lyase C, and **(c)** the β-roll domain of alkaline protease.

Adhesive proteins represent the largest group in the LRR superfamily. These proteins all contain similar 24-residue LRRs. One family of adhesive LRR-containing proteins includes the small proteoglycans: biglycan, fibromodulin, decorin, lumican, proteoglycan-Lb and osteoinductive factor (OIF, also called osteoglycin)[9]. Small proteoglycans bind various components of the extracellular matrix and growth factors (Table 1). Decorin and fibromodulin regulate collagen-fibril formation[9]; and OIF, in conjunction with the transforming growth factors TGF-β and TGF-β2, induces bone formation[14].

Several adhesive LRR-containing proteins are found in *Drosophila*. Toll, slit, connectin, chaoptin and flightless-I appear to orient cells during development; each protein is specific to particular tissues[15–19]. Connectin, chaoptin and a similar mammalian protein that may be involved in myelination (oligodendrocyte-myelin glycoprotein, OMGP)[20] are attached to the membrane via a glycosylphosphatidylinositol (GPI) anchor.

Another family of adhesive proteins comprises the proteins of the Ib–V–IX system of platelet glycoproteins. This complex constitutes the receptor for von Willebrand factor (vWF) and mediates the adhesion of platelets to injured vascular surfaces[21]. All polypeptides in the complex contain LRRs, but glycoproteins Ibβ and IX contain only one leucine-rich motif. Although it is unlikely that a single motif by itself would normally adopt the conformation observed in the crystal structure of RI, such a conformation could be stabilized by other proteins in the complex. Bacterial proteins YopM and IpaH interfere with platelet-mediated inflammatory responses, supposedly owing to homology with glycoprotein Ibα[22].

Signal-transducing receptors. The LRR superfamily contains several families of signal-transducing receptors. CD14 is a GPI-anchored receptor that, upon binding the complex between lipopolysaccharide and its binding protein, triggers intracellular protein tyrosine phosphorylation, which in turn leads to an antibacterial response by macrophages[23].

The proto-oncogene *trk* encodes a transmembrane protein, Trk, with a cytoplasmic tyrosine kinase domain[24]. The extracellular domain that contains LRRs is involved in high-affinity binding of nerve growth factor; other proteins from the Trk family selectively bind other neurotrophins (Table 1). TMK1, a plant protein with structural and biochemical properties similar to the receptor protein kinases, contains serine/threonine kinase in its cytoplasmic domain[25]. The receptor protein kinases are related to adhesive proteins such as Toll and glycoprotein Ib, as they are targeted to the plasma membrane and contain an extracellular LRR domain, a transmembrane domain and a carboxy-terminal cytoplasmic domain involved in signal transduction.

G-protein-coupled gonadotrophin receptors consist of an extracellular, hormone-binding LRR domain, a domain containing seven membrane-spanning segments and a short cytoplasmic tail[26]. Their ligands, chorionic gonadotrophin, luteinizing hormone, follicle-stimulating hormone and thyroid-stimulating

hormone, are heterodimeric glycoproteins composed of a specific β-subunit noncovalently attached to a common α-subunit.

Other signal-transducing proteins. Yeast adenylate cyclase contains LRRs in a domain that is important for the regulation of its catalytic activity by the RAS protein[27]. The trypanosomal leucine-rich protein may be related to this domain because it is encoded in the variant-surface-glycoprotein expression site adjacent to an adenylate cyclase[28].

Two LRR-containing proteins take part in the glucose repression pathway of yeast. GRR1 is the primary response element in this pathway[29], and CCR4 is required for the transcription of glucose-repressible alcohol dehydrogenase[30].

Other proteins. RAD1 and RAD7 constitute another functional group of LRR-containing proteins, as they are both involved in nucleotide excision repair in yeast. RAD1 forms a functional complex with RAD10, but the binding region lies outside the LRR-containing sequence[31]. Genetic evidence suggests that RAD7 may interact with RAD23 (Ref. 32).

RAD1 also plays a role in mitotic recombination[33]. Similarly, *DRT100*, a plant cDNA encoding a protein that contains LRRs and a nucleotide-binding motif, complemented deficiencies in homologous recombination in *Escherichia coli*[34].

The remaining LRR-containing proteins thus far appear unrelated to other members of the superfamily, but also have important functions: sds22 regulates protein phosphatase-1 function in fission-yeast mitosis[35]; the ribosome-binding protein p34 may function as the ribosome receptor in the secretory pathway[36]; the two LRR-containing subunits of the carboxypeptidase N tetramer stabilize the catalytic subunits and keep them in circulation[37]; and internalin confers invasive properties on *Listeria*[38].

Flanking cysteine clusters. Many LRR-containing proteins also contain homologous regions flanking the LRR domains (Fig. 1). These regions are characterized by four similarly spaced cysteines in a stretch of about 20 amino acids for the amino-flanking region and about 50 amino acids for the carboxy-flanking region[16,39]; the consensus sequences can be described as CP[~2x]CxC[~6x]C for the amino-flanking and PxxCxC[~20x]C[~20x]C for the carboxy-flanking regions (x denotes any amino acid). The amino-flanking cysteine clusters occur in all small proteoglycans, glycoproteins Ibβ and IX, Toll, slit, OMGP, CD14, the Trk family and gonadotrophin receptors; related sequences without all four cysteine residues conserved are present in leucine-rich α2-glycoprotein, glycoproteins Ibα and V, IpaH4.5, chaoptin and carboxypeptidase N. The carboxy-flanking cysteine clusters occur in glycoproteins Ibα, Ibβ, V and IX, Toll, slit and the Trk family; related sequences in which all four cysteines are not conserved are present in leucine-rich α2-glycoprotein and connectin. A different type of carboxy-flanking sequence containing two cysteines is found in small proteoglycans[39]. The flanking cysteine clusters appear to be a property of adhesive proteins and receptors.

LRRs with known function. In several cases, LRRs are responsible for specific functions. In RI, RNA1, small proteoglycans, chaoptin, CD14 and sds22, the tandem repeats comprise essentially the entire primary structure. In glycoprotein Ibα, the binding sites for vWF and thrombin have been mapped to the LRR-containing domain; although LRRs do not appear to bind vWF directly[21], point mutations in the LRR domain cause variants of Bernard-Soulier disease, a bleeding disorder (see, for example, Ref. 40). LRRs appear to be responsible for the regulation of yeast adenylate cyclase by RAS[27], for the hormone selectivity of the gonadotrophin receptors[41] and for the function of CCR4 (Ref. 30). The single leucine-rich motif in the type 1 human immunodeficiency virus Rev protein is the *trans*-activating region of the virus[42].

Concluding remarks

LRR-containing proteins represent an important superfamily involved in protein–protein interactions and signal transduction. The wide range of functions of LRR-containing proteins suggests that LRRs are generally useful protein-binding motifs; the variability of primary structures of these proteins suggests that these proteins do not necessarily share a common ancestor. The crystal structure of RI provides a model for the architecture of LRRs and helps us understand why LRRs may be useful for protein binding; the two structural features that might be responsible for the binding function are the solvent-exposed parallel β-sheet and the nonglobular shape of the structure. Owing to the variability of repeats and functional diversities of LRR-containing proteins, however, the crystal structure of RI can only represent a platform for further structural studies of these proteins.

Acknowledgements

We thank D. Crawford and B. Beutler for making us interested in ribonuclease inhibitor and leucine-rich repeats, and K. Fischer Lindahl and J. F. Sambrook for review of the manuscript and suggestions.

References

1 Takahashi, N., Takahashi, Y. and Putnam, F. W. (1985) *Proc. Natl Acad. Sci. USA* 82, 1906–1910
2 Kobe, B. and Deisenhofer, J. (1993) *Nature* 366, 751–756
3 Koo, Y. B., Slaughter, R. G. and Ji, T. H. (1991) *Endocrinology* 128, 2297–2308
4 Lee, F. S. and Vallee, B. L. (1993) *Prog. Nucleic Acid Res. Mol. Biol.* 44, 1–30
5 Krantz, D. D., Zidovetzki, R., Kagan, B. L. and Zipursky, S. L. (1991) *J. Biol. Chem.* 266, 16801–16807
6 Gay, N. J., Packman, L. C., Weldon, M. A. and Bama, J. C. J. (1991) *FEBS Lett.* 291, 87–91
7 Ohkura, H. and Yanagida, M. (1991) *Cell* 64, 149–157
8 Fresco, L. D., Harper, D. S. and Keene, J. D. (1991) *Mol. Cell. Biol.* 11, 1578–1589
9 Kresse, H., Hausser, H. and Schönherr, E. (1993) *Experientia* 49, 403–416
10 O'Shea, E. K., Klemm, J. D., Kim, P. S. and Alber, T. (1991) *Science* 254, 539–544
11 Yoder, M. D., Keen, N. T. and Jurnak, F. (1993) *Science* 260, 1503–1507
12 Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) *EMBO J.* 12, 3357–3364
13 Traglia, H. M., Atkinson, N. S. and Hopper, A. K. (1989) *Mol. Cell. Biol.* 9, 2989–2999
14 Madisen, L. *et al.* (1990) *DNA Cell Biol.* 9, 303–309
15 Hashimoto, C., Hudson, K. L. and Anderson, K. V. (1988) *Cell* 52, 269–279
16 Rothberg, J. M., Jacobs, J. R., Goodman, C. S. and Artavanis-Tsakonas, S. (1990) *Genes Dev.* 4, 2169–2187
17 Nose, A., Mahajan, V. B. and Goodman, C. S. (1992) *Cell* 70, 553–567
18 Krantz, D. E. and Zipursky, S. L. (1990) *EMBO J.* 9, 1969–1977
19 Campbell, H. D. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 11386–11390
20 Mikol, D. D., Gulcher, J. R. and Stefansson, K. (1990) *J. Cell Biol.* 110, 471–479
21 Roth, G. J. (1991) *Blood* 77, 5–19
22 Leung, K. Y. and Straley, S. C. (1989) *J. Bacteriol.* 171, 4623–4632
23 Weinstein, S. L., June, C. H. and DeFranco, A. L. (1993) *J. Immunol.* 151, 3829–3838
24 Meakin, S. O. and Shooter, E. M. (1992) *Trends Neurosci.* 15, 323–331
25 Chang, C. *et al.* (1992) *Plant Cell* 4, 1263–1271
26 Salesse, R. *et al.* (1991) *Biochimie* 73, 109–120
27 Suzuki, N. *et al.* (1990) *Proc. Natl Acad. Sci. USA* 87, 8711–8715
28 Ross, D. T. *et al.* (1990) *EMBO J.* 10, 2047–2053
29 Flick, J. S. and Johnston, M. (1991) *Mol. Cell. Biol.* 11, 5101–5112
30 Malvar, T., Biron, R. W., Kaback, D. B. and Denis, C. L. (1992) *Genetics* 132, 951–962

© 1994, Elsevier Science Ltd 0968–0004/94/$07.00

31 Siede, W., Friedberg, A. S. and Friedberg, E. C. (1993) *J. Bacteriol.* 175, 6345–6347
32 Perozzi, G. and Prakash, S. (1986) *Mol. Cell. Biol.* 6, 1497–1507
33 Schiestl, R. H. and Prakash, S. (1988) *Mol. Cell. Biol.* 8, 3619–3626
34 Pang, Q., Kays, J. B. and Rajagopal, I. (1992) *Proc. Natl Acad. Sci. USA* 89, 8073–8077
35 Stone, E. M., Yamano, H., Kinoshita, N. and Yanagida, M. (1993) *Curr. Biol.* 3, 13–26
36 Ohsumi, T. *et al.* (1993) *Biochem. J.* 294, 465–472
37 Tan, F. *et al.* (1990) *J. Biol. Chem.* 265, 13–19
38 Gaillard, J-L. *et al.* (1991) *Cell* 65, 1127–1141
39 Fischer, L. W. *et al.* (1991) *J. Biol. Chem.* 266, 14371–14377
40 Ware, J. *et al.* (1993) *J. Clin. Invest.* 92, 1213–1220
41 Braun, T., Schofield, P. R. and Sprengel, R. (1991) *EMBO J.* 10, 1885–1890
42 Malim, M. H., Bohnlein, S., Hauber, J. and Cullen, B. (1989) *Cell* 58, 205–214

Aquaporins: water channel proteins of plant and animal cells

Maarten J. Chrispeels and Peter Agre

Certain biological membranes, such as the erythrocyte plasma membrane, have a high osmotic water permeability, and such membranes have long been suspected of harboring water-channels. The molecular identity of these channels has now been established with the purification of water-channel proteins and the cloning of the genes encoding them. Homologous water-channel proteins, called 'aquaporins', are present in plants and animals. These channels are water selective and do not allow ions or metabolites to pass through them. Their discovery is providing new insights into how plant and animal cells facilitate and regulate the passage of water through their membranes.

CELLS ARE COMPOSED primarily of water, and movement of water into and out of cells is a key feature of numerous physiological functions in animals, plants and microorganisms. The plasma membranes of renal proximal tubules, certain other mammalian epithelia and erythrocytes are highly permeable to water, which moves in the direction of an osmotic gradient (see Refs 1, 2 for reviews). In plants, water moves through the living tissues in response to hydrostatic and osmotic forces. During the day, when the stomata are open, a constant transpiration stream moves through the tissues; water enters the roots and moves up the conductive tissues of the stem, and water vapor exits the leaves.

Numerous studies on the water permeability of mammalian epithelia and the hydraulic conductance of plant tissues and cells have been made, and the plasma membranes are thought to be the major barriers to water flow. The lipid bilayer is somewhat permeable to water resulting from simple diffusion, and the diffusional water permeability coefficient (P_d) of many membranes is <0.01 cm s^{-1}. This process cannot be inhibited pharmacologically and is characterized by a high Arrhenius activation energy ($E_a > 10$ kcal mol^{-1}, i.e. reduced water movement at lower temperatures where lipid packing is tighter). Pores or water channels were hypothesized to explain the more rapid transmembrane water movements characteristic of erythrocytes, certain mammalian epithelia and plant tissues. Unlike diffusional water permeability, channel-mediated water movement across erythrocyte membranes is reversibly inhibited by mercuric chloride and exhibits a low Arrhenius activation energy ($E_a < 5$ kcal mol^{-3}). The osmotic water permeability coefficient (P_f) of erythrocyte membranes measures approximately 0.02 cm sec^{-1}. Although erythrocyte membranes and certain other biological membranes have $P_f >> P_d$, suggesting the existence of water channels, the molecular identity of the water channels remained elusive until the recent identification of CHIP and a growing list of evolutionarily related water-channel proteins in animals and plants.

M. J. Chrispeels is at the Department of Biology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0116, USA; and **P. Agre** is at the Departments of Biological Chemistry and Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205-2185, USA.

© 1994, Elsevier Science Ltd 0968–0004/94/$07.00

An ancient family of channel proteins

The aquaporins are a functionally defined group of water-transport proteins that are part of an ancient family of channel proteins called the MIP family, after the first identified member, major intrinsic protein of the mammalian lens. MIP comprises approximately 50% of the total membrane protein of lens-fiber cells, and the cDNA encoding it was isolated from a bovine lens library[4]. Hydropathy analysis of the deduced amino acid sequence predicted a 263-residue integral membrane protein with cytoplasmic amino and carboxyl termini and six bilayer-spanning domains. The amino- and carboxy-terminal halves of MIP are themselves sequence related, forming a unique structure in which the topology of the first and second halves of the molecule are oriented at 180° to each other (Fig. 1); the MIP family already has more than 20 members, including: NOD26, a protein in the symbiosome membrane of soybean root nodules; the glycerol facilitator of *Escherichia coli* and other bacteria; a suppressor of a *Saccharomyces cerevisiae* growth defect on fermentable sugars; a developmentally regulated *Drosophila* brain protein; and multiple tonoplast intrinsic proteins in plants (Ref. 5 and references therein). The functions of many MIP family members have yet to be established, but several plant and animal MIP-family members have been shown to be water-specific membrane channels.

Structure of the aquaporins

As with MIP, the 28 kDa channel-forming integral protein (CHIP) was first

EDITOR
Jo McEntyre

ASSISTANT EDITOR
Catherine Brooksbank

EDITORIAL SECRETARY
Christine Oliver

PRODUCTION
Glyn Williamson

PUBLISHER
Peter Desmond

EDITORIAL BOARD

Editor in Chief
Tim Hunt
ICRF Clare Hall Laboratories, UK

M. Buckingham
Institut Pasteur, Paris, France

A. R. Fersht
University of Cambridge, UK

M-J. Gething
HHMI, University of Texas, Dallas, USA

C. Guthrie
University of California, San Francisco, USA

N. Heintz
HHMI, Rockefeller University, New York, USA

C-H. Heldin
Ludwig Institute for Cancer Research, Uppsala, Sweden

W. Junge
University of Osnabrück, Germany

A. Lamond
EMBL, Heidelberg, Germany

P. Schimmel
MIT, Cambridge, USA

P. Srere
VA Medical Center, Dallas, USA

P. H. von Hippel
University of Oregon, Eugene, USA

J. Witkowski
Banbury Center, Cold Spring Harbor Laboratory, New York, USA

M. Yanagida
Kyoto University, Japan

ADVERTISING
Phil Abrahams
(Business Development)

Thelma Reid
(Sales)

PUBLISHING OFFICE
Elsevier Trends Journals,
68 Hills Road, Cambridge,
UK CB2 1LA

Tel: [44] 223 315961 (all depts)
Fax: [44] 223 464430 (ed.)
Tel: [44] 223 311114 (ads)
Fax: [44] 223 321410 (ads)

October 1994 (226) Volume 19 No. 10 pp. 395–437

trends in BIOCHEMICAL SCIENCES
is an official publication of the INTERNATIONAL UNION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

ELSEVIER TRENDS JOURNALS

C O N T E N T S

HEALTH SCIENCES LIBRARY
University of Wisconsin
NOV 11 1994

COVER

Titin and nebulin are giant muscle proteins composed of multiple repeating domains. They are thought to act as 'protein rulers', stabilizing and regulating the assembly of actomyosin. See article on p. 405 for details.

Cover design by Rosie Wright.

SUBSCRIPTION INFORMATION

PERSONAL SUBSCRIPTION: (12 issues starting anytime in 1994)
£69.00 UK/Rest of World; $110.00 USA/Canada. Must be prepaid.

STUDENT SUBSCRIPTION: (12 issues starting anytime in 1994)
£34.00 UK/Rest of World; $55.00 USA/Canada. Proof of student status required.

LIBRARY/INSTITUTIONAL SUBSCRIPTION:
(Vol. 19 1994 12 issues plus hard-bound compendium) $490.00/£318.00 Worldwide

Prices include handling and air delivery worldwide. US Mailing Agent: USA Postmaster: second class postage paid by Newark, NJ, USA. Postmaster: send all USA address changes to *Trends in Biochemical Sciences*, c/o Virgin Mailing & Distribution, 10 Corporate Road, Irvington, New Jersey 07111, USA. See page XII for further copyright information.
USA and Canada Customer Service and Advertising Enquiries: Elsevier Trends Journals, Elsevier Science Inc., 660 White Plains Road, Tarrytown, NY 10591 5153, USA. Tel: 914 524 9200; Fax: 914 333 2444.
UK and Rest of World Subscription Enquiries: Elsevier Trends Journals, Elsevier Science Ltd, Journals Circulation Department, PO Box 800, Kidlington, Oxford, UK OX5 1DX. Tel: [44] 865 843300; Fax: [44] 865 843940

Indexed/Abstracted in: ABSTRACTS IN BIOCOMMERCE, ABSTRACTS ON HYGIENE, BIOLOGICAL AGRICULTURE, BIOCULTURE, BIOLOGICAL ABSTRACTS, BIOSCIENCE INFORMATION SERVICE, BIOTECH ABSTRACTS, CABS, CAMBRIDGE SCIENTIFIC ABSTRACTS, CHEMICAL ABSTRACTS, CURRENT BIOTECHNOLOGY ABSTRACTS, CURRENT CONTENTS (LS), ENERGY INFORMATION ABSTRACTS, ENVIRONMENT ABSTRACTS, EXCERPTA MEDICA, INDEX TO SCIENTIFIC REVIEWS, MEDLARS, SCIENCE CITATION INDEX.

ABC AUDIT BUREAU OF CIRCULATIONS
BUSINESS PRESS

PRINTED IN THE UK.

ISSN 0968-0004

JOURNAL NUMBER: 02115

TU BMB

8–11 November 1994

Polymerase Chain Reaction (PCR) Applications/Cycle DNA Sequencing, Maryland, USA. (ATCC/Workshop Manager, 12301 Parklawn Drive, Rockville, MD 20852, USA. Tel: +1 301 231 5566. Fax: +1 301 770 1805.)

8–21 November 1994

Molecular Markers for Plant Breeding and Plant Genetics – Course, Cold Spring Harbor, NY, USA. (Meetings Office, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724-2213, USA.)

8–21 November 1994

Monoclonal Antibodies from Combinatorial Libraries – Course, Cold Spring Harbor, NY, USA. (Meetings Office, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724-2213, USA.)

14–18 November 1994

Ultrastructural Immunocytochemistry Course, Sutton, UK. (Royal Microscopical Society, 37–38 St Clements, Oxford, UK OX4 1AJ.)

15–18 November 1994 ★

Biopolymers and Bioproducts: Structure, Function and Applications, Bangkok, Thailand. (Prof. Jisnuson Svasti, Department of Biochemistry, Faculty of Science, Mahidol University, Rama VI Road, Bangkok 10400, Thailand. Tel: +66 2 2461360 to 69, ext. 4305/4306. Fax: +66 2 2480375 or 2596634.)

16–18 November 1994

Cell Culture and Hybridomas: Quality Control and Cryopreservation Techniques, Maryland, USA. (ATCC/Workshop Manager, 12301 Parklawn Drive, Rockville, MD 20852, USA. Tel: +1 301 231 5566. Fax: +1 301 770 1805.)

16–20 November 1994 ★

Immunobiology of Proteins and Peptides, Rio Rico, AZ, USA. (Dr Garvin S. Bixler, Jr, Univax Biologics Inc., 12280 Wilkins Avenue, Rockville, MD 20852, USA. Tel: +1 301 770 3099. Fax: +1 301 770 2014.)

20–22 November 1994 ★

International Symposium on Genomic Imprinting, Florence, Italy. (Prof. M. L. Giovannucci Uzielli, Human Genetics Service, University of Florence, Via Masaccio 209, Firenze 50132, Italy. Tel: +39 55 5662 931. Fax: +39 55 5662 916.)

28–29 November 1994

Adhesion Proteins, Paris, France. (Prof. B. Boris Vargaftig, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France. Tel: +33 1 4568 8682/8688. Fax: +33 1 4568 8703.)

1–2 December 1994

Cycling out of Control: the Cell Cycle in Growth, Differentiation and Disease, Paris, France. (Pippa Burnett, *Nature*, Macmillan Magazines Ltd, 4 Little Essex Street, London, UK WC2R 3LF. Tel: +44 (0)71 836 663. Fax: +44 (0) 71 279 5417.)

7–9 December 1994

IBC: 5th Annual International Conference, La Jolla, CA, USA. (IBC USA Conferences, 225 Turnpike Road, Southborough, MA 01772-1647, USA. Tel: +1 508 481 6400. Fax: +1 508 481 7911.)

12–14 December 1994 ★

Molecular Biology and Cancer Network: 11th Annual Meeting, Warwick, UK. (K. Weston, Institute of Cancer Research, Fulham Road, London, UK SW3 6JB. Fax: +44 (0)71 352 3299.)

14–16 December 1994

An Introduction to Genetic Engineering, Coventry, UK. (Carol Howes, Dept of Biological Sciences, University of Warwick, UK CV4 7AL. Tel: +44 203 523517. Fax:+44 203 523568.)

15 December 1994 ★

Cell Cycle Control and Apoptosis: Basic Mechanisms and Implications for Treating Malignant Disease, London, UK. (Mrs J. Wenley, Symposium Administrator, Department of Molecular Pathology, UCL Medical School, Windeyer Building, 46 Cleveland Street, London, UK W1P 6DB. Tel: + 071 380 9343. Fax: +44 71 387 3310.)

5–11 January 1995

Oncogenes: 20 Years Later, Keystone, CO, USA. (Keystone Symposia, Drawer 1630, Silverthorne, CO 80498, USA. Tel: +1 303 262 1230.)

26–28 January 1995 ★

Ions and Drugs: Active Transport in Eukaryotic Cells, Villard de Lans, France. (Dr Yves Dupont, DBMS/BMC, Grenoble, France. Tel: +33 76 88 46. Fax: +33 76 88 54 87.)

5–9 February 1995 ★

Gordon Research Conference on Carotenoids, Ventura, CA, USA. (James A. Olson, Iowa State University, Biochemistry and Biophysics, Ames, IA 50011, USA. Tel: +1 515 294 3068. Fax: +1 515 294 4141.)

6–7 April 1995 ★

BSCB/British Connective Tissue Society – Molecular and Cell Interactions in Connective Tissue, Manchester, UK. (Karl Kadler and Martin Humphries, University of Cambridge, Tennis Court Road, Cambridge, UK CB2 1QP. Fax: + 44 61 275 5082.)

18–22 April 1995 ★

BSCB/BSDB Spring Meeting – Stem Cells, Canterbury, UK. (Enrico Coen and Fiona Watt, CRC/Wellcome, Cambridge, UK.)

18–22 April 1995 ★

BSCB Spring Meeting – Organization of the Cell Nucleus, Canterbury, UK. (Ron Laskey, Tony Kouzarides and Colin Dingwall, CRC/Wellcome, Cambridge, UK.)

18–22 April 1995 ★

BSCB Spring Meeting – Workshop on Muscle Cells, Canterbury, UK. (Simon Hughes and Bob Simmons, BSCB/BSDB, University of Kent. Fax: +44 227 763912.)

18–22 April 1995 ★

BSCB Spring Meeting – Cell Movements in Development, Canterbury, UK. (Anthony Baines (BSCB) and Clive Wilson (BSDB). Fax: +44 227 763912.)

22–24 May 1995 ★

22nd Stadler Genetics Symposium: All Aspects of Genome Analysis Including Plant and Animal, Columbia, MO, USA. (J. P. Gustafson, 206 Curtis Hall, University of Missouri-Columbia, Columbia, MO 65211, USA.)

18–23 August 1996 ★

9th International Congress of Human Genetics, Rio de Janeiro, Brazil. (Dr Henrique Krieger, Secretary General, 9th International Congress of Human Genetics, Instituto de Ciências Biomédicas, Universidade de São Paulo, Cx. Postal 66208, 05389-970 São Paulo, SP-Brazil. Tel: +55 11 818 7328. Fax: +55 11 818 7417 or +55 11 210 4817. Email: 9ichg@biomed.icb2.usp.br)

Note: *Trends in Biochemical Sciences* is apolitical. The views and opinions expressed by its writers and correspondents do not necessarily reflect those of the editorial board, staff editor, or the publishers. No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Although all advertising material is expected to conform to ethical (medical) standards, inclusion in this publication does not constitute a guarantee or endorsement of the quality or values of such product or the claims made of it by the manufacturer.

© 1994 Elsevier Science Ltd. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical photocopying, recording or otherwise without the prior permission of the publisher, Elsevier Trends Journals.

Special Regulations for readers in the USA. This journal has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal use of specific clients. This consent is given on the condition that the copier pays through the center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by Section 107 or 108 of the US Copyright Law. The appropriate fee should be forwarded with a copy of the first page of the article to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. If no code appears in an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1981 may be copied for a per-copy fee of US$7.00, also payable through the Center. This consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotion purposes, or for creating new collective works. Special written permission must be obtained from the publisher for such copying.

Special regulations for authors. Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright of the article to the publisher. This transfer will ensure the widest possible dissemination of information.

The copyright clearance center fee code for *Trends in Biochemical Sciences* is 0968-0004/$07.00.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.